IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Patent Application of: **FIENBLIT et al.**
Serial No.: 10/673,529
Filed: September 29, 2003

For: ASYNCHRONOUS DATA MIRRORING WITH LOOK-AHEAD SYNCHRONIZATION RECORD

Group Art Unit: 2162
Examiner: Giovanna B. Colan

RULE 132 DECLARATION OF MICHAEL FACTOR

I, the undersigned, Michael Factor, of 32 Rachel Street, Haifa, Israel, hereby declare as follows:

1. I have worked in the field of data storage systems for the past 7 years and have 105 academic publications, patents and patent applications on data storage systems and related subjects. I received a Ph.D. in 1990 from Yale University, in New Haven, CT, USA. Since 1991, I have worked at IBM Haifa Research Laboratory in Haifa, Israel, working as a Storage Architect and Researcher (2001-2007) and IBM Distinguished Engineer (since 2006), in the field of data storage systems.

2. I have read and familiarized myself with U.S. Patent Application 10/673,529 (to which I refer hereinafter simply as "the Application") and with the claims in the Application as they are to be amended in a response to the outstanding Official Action in this case. The claims cover a data storage system, as well as methods and software for managing such a system. The data storage system includes primary and secondary storage subsystems,

each including non-volatile storage media, such as disk drives. When the system receives data from a host processor, it stores the data in the specified location in the primary subsystem and copies the data to the secondary subsystem for storage there, as well, to create a mirror of the primary subsystem. The system uses a novel method of predictive record-keeping to maintain synchronization between the primary and secondary subsystems to enable recovery from failure and to restore data from the secondary to the primary subsystem in case of failure.

3. The purpose of the predictive record that is recited in the claims, in other words, is to provide a basis for deciding which data to restore from the secondary subsystem to the primary subsystem after a failure of the primary subsystem. Other storage systems with which I am familiar use meticulous record-keeping of every host write operation within a grain (i.e., the unit of tracking update such as a 64K track) for this purpose without predicting writes to other grains. The record must be maintained in non-volatile memory to avoid loss of the information in case of failure. Such frequent access to non-volatile storage either slows the performance of the system or typically requires the use of large amounts of costly high-speed non-volatile memory to store the record. The predictive record recited in the claims of the Application alleviates this problem, since it permits the primary subsystem to record multiple locations in the record at once, whereby most write operations do not require the record to be updated (as

explained in paragraph 0011 in the Application.)

4. The structure and operation of the claimed data storage system (with the exception of the use of the novel synchronization record) resemble those of large-scale storage systems produced by IBM and its competitors in this area, such as EMC. One example of such a system is the IBM Enterprise Storage Server (ESS), mentioned in paragraph 0004 in the Application. As explained in the Application, each of the primary and secondary subsystems is a self-contained unit, with its own control unit (CU), cache and non-volatile storage, which is typically a collection of disks or disk arrays (see Fig. 1 and paragraph 0038, for example).

5. Storage systems of this type are typically assembled, installed and maintained by application engineers and customer service engineers. Such engineers generally have a bachelor's degree in computer science or engineering. Their job is to choose the appropriate hardware components, such as control units, cache memory and disk drives; to set up the controller software according to the hardware configuration; to install the necessary software on host computers for interacting with the storage system; and to troubleshoot problems in system configuration and operation. These are the characteristics and skills of persons having ordinary skill in the art relevant to data storage systems of the type recited in the claims of the Application.

6. Such persons of ordinary skill do not deal with the internal workings of the hardware components in the data

storage system. For example, when there is a failure in a hardware component, the engineer will typically simply replace it, rather than attempting to fix it. Persons of ordinary skill in storage systems (and even experts in the field, such as myself) generally regard the hardware components, such as disks and disk arrays, as interchangeable "black boxes" and do not attempt to understand or modify the internal workings of such components.

7. In the Official Action dated September 24, 2007, the Examiner rejected all of the claims in the Application for obviousness over Black (U.S. Patent 6,978,324) in view of Willis et al. (U.S. Patent 5,550,998) and further in view of Bakke et al. (U.S. Patent Application Publication 2003/0023808). Having reviewed the Official Action and the cited references against the language of the amended claims, I believe this rejection should be withdrawn for a number of reasons, which I will detail below.

8. Black describes controlling read and write accesses to a logical entity, in which a host computer uses an "enterprise logical volume identifier," or "ELVID." In rejecting the claims in the present Official Action, the Examiner cited col. 27, lines 47-50 and 59-61, in Black, which refer Black's host table 181. According to the cited passage, the host computer uses the host table "to identify an ELVID with a particular physical address identifier specified in the data access request by the host." The logical volume identifier is independent of the actual location of storage of the data in the logical

volume, and the same logical volume identifier may be maintained when the logical volume is moved among storage elements (col. 22, lines 50-53). Furthermore, the ELVID, and hence the content of the table, are independent of the volume content, and so are unrelated to ongoing write operations. As such, this function is irrelevant to problems of data mirroring and failure recovery within the storage system that are addressed by the claims in the Application, which are concerned with which subset of the volume data has to be copied.

9. Bakke describes a method for maintaining data coherency in a dual I/O adapter, which includes primary and secondary adapters (abstract). Bakke defines such I/O adapters as "interfaces between a computing system and a storage subsystem" (paragraph 0007), i.e., elements that could interface between the host computer and storage subsystems that are shown in Fig. 1 of the Application, for example. In other words, Bakke's adapter likewise stands outside the storage system itself. This relationship is shown clearly in Bakke's Fig. 2, in which two of his I/O adapters are connected to a single storage device (DASD 250).

10. Willis describes a pre-seeking process for enhancing operation of a mass storage system. Willis, however, uses the term "mass storage system" in a way that is different from the storage systems in Black, Bakke and the Application. Willis defines a "mass storage system" as a group of individual disk drives that may be used in place of a "single disk mass storage system" (col. 1, lines 19-27). All of the embodiments he describes are

based on RAID (redundant array of independent disks) systems. In his example of a RAID system in Fig. 1, Willis shows a single RAID mass storage system connected to a host computer through a host bus (see description in col. 4, lines 20-22). The array controller in the RAID system uses a mapping table to associate I/O commands from the host computer with the locations on the disks where the data are to be written or read (col. 5, lines 5-7, cited by the Examiner).

11. Willis is most concerned with the inner workings of the disk drives in the array, and specifically with pre-seeking or pre-locating the transducer (which he calls a read/write head) of the disk in order to reduce the amount of time necessary to seek the transducer to the appropriate track for a desired read or write operation (col. 6, lines 40-45). For this purpose, Willis uses the predictability of the write algorithm to pre-seek the transducer to the location at which the next write operation is anticipated to be performed (col. 2, line 64 - col. 3, line 4). This write algorithm is used solely for internal purposes within the disk array, in order to minimize the seek time of the transducer (abstract). Willis, in other words, uses prediction to solve the problem of manipulating the transducer of an individual disk or disk array. There is no suggestion in Willis that such prediction could be used, for example, in mirroring data from one disk to another in a RAID system.

12. A single disk or disk array of the sort described by Willis could serve as storage media 34 in the storage subsystems that are described in the Application (as

noted in paragraph 0038 of the Application). As I explained above, however, the person having ordinary skill in the art relevant to the claims in the Application would regard Willis's disk array as a "black box," and would not be concerned with its inner workings. This person of ordinary skill would not have looked to Willis's method of manipulating a transducer within the disk array as a solution to the problems of data mirroring and failure recovery among different storage subsystems that are addressed by the claims in the Application.

13. Furthermore, as I have pointed out above, Black is concerned with functions carried out by the host computer, external to the storage system. Therefore, the person of ordinary skill would have seen no point in incorporating Willis's method of prediction into Black's system. Such incorporation could not have led to any sort of method of predictive record-keeping of write operations that could be used in the sort of data mirroring and failure recovery that is recited in the claims of the Application.

14. To summarize, for the reasons that I have explained above, I believe the method, apparatus and software product recited in the claims of the Application to be novel and non-obvious over the cited references, as well as over all other prior art of which I am aware.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and conjecture are thought to be true; and further that these statements were made with the

knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application of any patent issued thereon.

Michael Factor, Citizen of USA and Israel
32 Rachel Street, Haifa, Israel
December 20, 2007